Stochastic Better-Reply Dynamics in Games

By Jens Josephson*

July 3, 2000

Abstract. In Young (1993, 1998) agents are recurrently matched to play a finite game and almost always play a myopic best reply to a frequency distribution based on a sample from the recent history of play. He proves that in a generic class of finite n-player games, as the mutation rate tends to zero, only strategies in certain minimal sets closed under best replies will be played with positive probability. In this paper we alter Young's behavioral assumption and allow agents to choose not only best replies, but also better replies. The better-reply correspondence maps distributions over the player's own and her opponents' strategies to those pure strategies which gives the player at least the same expected payoff against the distribution of her opponents' strategies. We prove that in *all* finite n-player games, the limiting distribution will put positive probability only on strategies in certain minimal sets closed under better replies. This result is consistent with and extends Ritzberger's and Weibull's (1995) results on the equivalence of asymptotically stable strategy-sets and closed sets under better replies in a deterministic continuous-time model with sign-preserving selection dynamics.

Keywords: Evolutionary game theory, Markov chain, stochastic stability, better replies

JEL classification: C72, C73

1. Introduction

Young (1993, 1998) develops a model where agents from each of n finite populations every period are randomly chosen to play a finite n-player game. Each of the agents forms beliefs of her opponents' play by inspecting an independent sample of recent

*I am grateful for helpful comments from Jörgen Weibull, Martin Dufwenberg and Alexander Matros. I also want to thank the Jan Wallander and Tom Hedelius Foundation for financial support.

strategy realizations from a finite history of past play and calculating the corresponding frequency distributions over her opponents' pure strategies. She thereafter chooses a pure strategy which is a best reply given her beliefs. With a small probability the agents instead experiment or make mistakes and play any strategy at random. This results in an ergodic Markov chain with a unique stationary distribution on the space of histories.

In the present paper we alter Young's behavioral assumption by allowing agents to play not only best replies but also better replies given their beliefs. Formally, the *better-reply correspondence* maps distributions over the player's own and her opponents' strategies to those pure strategies which gives the player at least the same expected payoff against the distribution of her opponents' strategies. The best-reply and the better-reply correspondences have important properties in common. Among other things, they both belong to the class of *behavior correspondences* (Ritzberger and Weibull, 1995). This class contains upper hemi-continuous correspondences such that the image of any product distribution includes the set of best-replies. The similarities between the two correspondences imply that Young's (1998) predictions in weakly acyclic games and in 2×2 coordination games are unaffected if best replies are replaced by better replies.

A product set of pure strategies is *closed* under the better-reply correspondence if the image under this correspondence of every distribution with support in the set is contained in the set. It is a *minimal closed set* if it does not properly contain a subset with this property. Such sets are important in our analysis of the stochastic adaptive model where agents apply the better-reply correspondence. We show that in finite n-player games, as the experimentation probability tends to zero, only strategies in certain minimal closed sets under better replies are played with positive probability. This is analogous to the result in Young (1998) for the best-reply model. He proves that in a generic class of finite games, only strategies in certain minimal closed sets under best replies will be played with positive probability in the limit.[1] Due to a specific property of the better-reply correspondence, our result is valid for all games, whereas Young have to assume the game is nondegenerate in a specific sense.

Our result is also consistent with and extends previous findings on deterministic dynamics in a continuous-time setting. Ritzberger and Weibull (1995) prove equivalence between asymptotically stable sets under sign-preserving selection dynamics and closed sets under better-replies. Hence our model provides a way to select among such asymptotically stable sets.

This paper is organized as follows. In Section 2 we present the stochastic process and introduce basic definitions. Section 3 contains results on the convergence of the stochastic process. In Section 4 we relate our findings to previous results on set-wise stability under regular selection dynamics. Section 5 contains examples and Section 6 discusses extensions of the model. The proof of the main theorem can be found in the Appendix.

2. Model

The basic setting is similar to that of Young (1993, 1998). Let Γ be a finite n-player game in strategic form. Let X_i be the finite set of pure strategies x_i available to player $i \in \{1, ..., n\} = N$ and let $\Delta(X_i)$ be the set of probability distributions p_i

[1]Another paper with a similar result is Josephson and Matros (2000) who also use Young's (1993) stochastic setup, but assume individuals are imitators who choose the strategy with the highest average payoff in their sample. Under this behavioral assumption, for sufficiently small mutation rates, strategies in certain *minimal closed sets under the better-reply graph* will be played with arbitrary high probability. These sets are closely related to minimal closed sets under better replies, but in other respects the dynamics and predictions in the model are quite distinct from those in the present paper.

over these strategies. Define the product sets $X = \prod X_i$ and $\square(X) = \prod_i \Delta(X_i)$ with typical elements x and p respectively. Let $p_i(x_i)$ denote the probability mass on strategy x_i and let $p(x) = \prod_{i \in N} p_i(x_i)$. The notation $x_{-i} \in \prod_{j \neq i} X_j$ and analogously $p_{-i} \in \prod_{j \neq i} \Delta(X_j) = \square(X_{-i})$ is used to represent the strategies and distributions of strategies of players other than i. Note that when we write "strategies" and "strategy-tuples" we always refer to pure strategies and pure strategy-tuples. Let $C_1, ..., C_n$ be n finite and non-empty populations of agents. Each member of population C_i is a candidate to play role i in the game G and has payoffs represented by the utility function $\pi_i : X \rightarrow R$, and expected payoffs represented by the function $u_i : \square(X) \rightarrow R$. In slight abuse of notation we write $u_i(x_i, p_{-i})$ instead of $u_i(p_i, p_{-i})$ if $p_i(x_i) = 1$.

The following two correspondences are instrumental in the subsequent analysis.

Definition 1. Let $\beta = \prod_i \beta_i : \square(X_{-i}) \twoheadrightarrow X$ be the *best-reply correspondence*, defined by $\beta_i(p_{-i}) = \{x_i \in X_i \mid u_i(x_i, p_{-i}) - u_i(x_i', p_{-i}) \geq 0 \ \forall x_i' \in X_i\}$.

Definition 2. (Ritzberger and Weibull, 1995) Let $\gamma = \prod \gamma_i : \square(X) \twoheadrightarrow X$ be the *better-reply correspondence*, defined by $\gamma_i(p) = \{x_i \in X_i \mid u_i(x_i, p_{-i}) - u_i(p) \geq 0\}$.

In other words γ_i assigns to each product distribution $p \in \square(X)$ those pure strategies x_i which give i at least the same expected payoff as p_i.[2] Note that the better-reply correspondence is defined as a mapping from the product of the simplicies for *all* players, whereas the best-reply correspondence is a mapping from the product of only the opponents' simplicies.[3]

[2]It is evident that γ_i is u.h.c. and $\beta_i(p) \subset \gamma_i(p)$ for all players and product distributions. As will be discussed below this implies that γ is a *behavior correspondence* (Ritzberger and Weibull, 1995).

[3]The best-reply correspondence can of course also be represented as a mapping from the set of product distributions to the set of pure strategy-tuples $\beta = \prod_i \beta_i : \square(X) \twoheadrightarrow X$ with $\beta_i(p) = \beta_i(p_i', p_{-i})$ for all $p_i' \in \Delta(X_i)$.

Let $t = 1, 2, ...$ denote successive time periods. The stage game Γ is played once each period. In period t, one individual is drawn at random from each of the n populations and assigned to play the corresponding role. The individual in player position i chooses a pure strategy x_i^t from his strategy space X_i according to a rule that will be defined below. The strategy-tuple $x^t = (x_1^t, ..., x_n^t)$ is recorded and referred to as *play* at time t. The *history* or *state* at time t is the sequence $h^t = (x^{t-m+1}, ..., x^t)$ where m denotes the *memory size* of all individuals. Let $H = X^m$ be the finite set histories of length m and let h be an arbitrary element of this set.

Strategies are chosen as follows. Assume an arbitrary initial history $h^m = (x^1, ..., x^m)$ at time m. In every subsequent period each individual drawn to play the game inspects s plays drawn without replacement from the most recent m periods. The draws are independent for the various individuals and across time. For each $x_i \in X_i$, let $p_i(x_i \mid h)$ be the conditional probability that agent i chooses strategy x_i given history h. We assume that $p_i(x_i \mid h)$ is independent of t and that $p_i(x_i \mid h) > 0$ if and only if there exists a sample of size s from the history h, consisting of n independent draws and with a sample distribution of $\hat{p} \in \Box(X)$, such that $x_i \in \gamma_i(\hat{p})$, where γ_i is the better-reply correspondence defined above. Unlike in Young (1993, 1998), the individuals here play a better reply and not a best reply to this sample distribution.

Given a history $h^t = (x^{t-m+1}, ..., x^t)$ at time t, the process moves in the next period to a state of the form $h^{t+1} = (x^{t-m+2}, ..., x^t, x^{t+1})$. Such a state is called a *successor* of h^t. Our behavioral assumptions imply that the process moves from a current state h to a successor state h' in each period according to the following transition rule. If x is the rightmost element of h, the probability of moving from h to h' is $P_{hh'}^{\gamma,m,s,0} = \prod_{i=1}^n p_i(x_i \mid h)$ if h' is a successor of h and 0 if h' is not a successor of h. This defines a finite Markov chain on the finite state space of histories H. We call the process $P^{\gamma,m,s,0}$ *γ-adaptive play with memory m and sample size s.* We generally refer to it as the unperturbed process. A *recurrent class* E_k of the process $P^{\gamma,m,s,0}$ is a set of states such that there is zero probability of moving from any state in the class to any state outside, and there is a positive probability of moving from any state in the class to any other state in the class. A state h is *absorbing* if it constitutes a singleton recurrent class.

We also define a perturbed process in a manner completely analogous to Young (1993, 1998). Formally, in each period there is some small probability ε, that player i experiments by choosing a strategy at random from X_i instead of according to the better-reply correspondence. The event that i experiments is assumed to be independent of the event that j experiments for every $j \neq i$. For every i let $q_i(x_i \mid h)$ be the conditional probability that i chooses $x_i \in X_i$, given that i experiments and the process is in state h. We assume that $q_i(x_i \mid h)$ is independent of t and that $q_i(x_i \mid h) > 0$ for all $x_i \in X_i$. Suppose that the process is in state h at time t. Let J be a subset of j players. The probability is $\varepsilon^j(1-\varepsilon)^{n-j}$ that exactly the players in J experiment and the others do not. Conditional on this event the transition probability of moving from h to h' is $Q^J_{hh'} = \prod_{i \in J} q_i(x_i \mid h) \prod_{i \notin J} p_i(x_i \mid h)$ if h' is a successor of h and x is the rightmost element of h' and 0 if h' is not a successor of h. This gives the following transition probability of the perturbed Markov process: $P^{\gamma,m,s,\varepsilon}_{hh'} = (1-\varepsilon)^n P^{\gamma,m,s,0}_{hh'} + \sum_{J \subset N, J \neq \emptyset} \varepsilon^{|J|}(1-\varepsilon)^{n-|J|} Q^J_{hh'}$. We call the process $P^{\gamma,m,s,\varepsilon}$ *γ-adaptive play with memory m, sample size s, experimentation probability ε and experimentation distributions q_i.*

This process is irreducible and thus has a unique stationary distribution μ^ε. We study this distribution as ε tends to zero. In our analysis we use the following standard definitions. A state h is *stochastically stable* if $\lim_{\varepsilon \to 0} \mu^\varepsilon(h) > 0$. For each pair of distinct recurrent classes E_k and E_l, a *kl-path* is a sequence of states $\zeta = (h^1, h^2..., h^q)$ that begins in E_k and ends in E_l. The *resistance* of this path is the sum of the resistances on the edges that compose it. Let r_{kl} be the least resistance over all kl-paths. Construct a complete directed graph with one vertex for each recurrent class. The weights on the directed edge $E_k \to E_l$ is r_{kl}. A *tree* rooted at E_l is a set directed edges such that, from every vertex different from E_l, there is a unique directed path in the tree to E_l. The *resistance* of a rooted tree is the sum of the resistances on the edges that compose it. The *stochastic potential* $\rho(E_l)$ of a recurrent class E_l is the minimum resistance over all trees rooted at E_l.

3. Results

In this section we investigate the properties of the limiting stationary distribution of γ-adaptive play as the experimentation probability tends to zero.[4] In what follows we will refer to the stochastic process in Young (1993, 1998) as β-adaptive play in order to distinguish it from γ-adaptive play. We will first prove our main result, that in finite n-player games the limiting distribution of γ-adaptive play puts positive probability only on histories with support in certain minimal closed sets under better replies. This result is similar to Theorem 7.2 in Young (1998) for β-adaptive play, but holds more generally. Thereafter we will show that two other results in Young hold also for γ-adaptive play.

3.1. Stochastic Stability of Minimal Closed Sets under Better-Replies.

In order to state our main result some further definitions are needed. Let $\mathcal{X}$ be the collection of all nonempty product sets $Y \subset X$. Let $\Delta(Y_i)$ be the set of probability distribution with support in Y_i and let $\square(Y) = \prod_{i \in N} \Delta(Y_i)$ be the corresponding product set.

Definition 3. (Ritzberger and Weibull, 1995) A set $Y \in \mathcal{X}$ is *closed under better replies* if $\gamma(\square(Y)) \subseteq Y$. A set $Y \in \mathcal{X}$ is a *minimal closed set under better replies* if

[4] This limiting distribution exist by Lemma 2 and Theorem 5 in the Appendix.

it is closed under better replies and contains no proper subset with this property.

Important to our analysis is the following lemma, which implies that our main theorem applies to *any* finite game and not only to a generic class of games as the corresponding theorem in Young (1998). Define $\tilde{\gamma}$ as the better-reply correspondence with pure strategy domain. That is $\tilde{\gamma} = \prod \tilde{\gamma}_i : X \twoheadrightarrow X$, where
$\tilde{\gamma}_i(x) = \{x'_i \in X_i \mid \pi_i(x'_i, x_{-i}) - \pi_i(x) \geq 0\}$.

Lemma 1. *If $x_i \in \gamma_i(p)$, then there exists a pure strategy-tuple $y \in X$ with $p(y) > 0$ such that $x_i \in \tilde{\gamma}_i(y)$.*

Proof: This proof uses the multilinearity of the function u_i. Consider an arbitrary distribution $p \in \Box(X)$, and a pure strategy $x_i \in X_i$ such that $x_i \in \gamma_i(p)$. We will first show that $\exists y_i \in X_i$ with $p_i(y_i) > 0$ such that $u_i(x_i, p_{-i}) - u_i(y_i, p_{-i}) \geq 0$. By definition of the better-reply correspondence we have that $u_i(x_i, p_{-i}) - u_i(p) \geq 0$. Write this difference in the following way:

$$u_i(x_i, p_{-i}) - u_i(p) = \sum_{y_i \in X_i} p_i(y_i) \left[u_i(x_i, p_{-i}) - u_i(y_i, p_{-i})\right]. \tag{1}$$

Clearly, if the left-hand side is non-negative, at least one of the terms in the sum on the right-hand side with $p_i(y_i) > 0$ must be non-negative. Hence $\exists y_i \in X_i$ with $p_i(y_i) > 0$ such that $u_i(x_i, p_{-i}) - u_i(y_i, p_{-i}) \geq 0$.

Second, write the last difference in the following way:

$$u_i(x_i, p_{-i}) - u_i(y_i, p_{-i}) = \sum_{y_{-i} \in X_{-i}} p_{-i}(y_{-i}) \left[\pi_i(x_i, y_{-i}) - \pi_i(y_i, y_{-i})\right]. \tag{2}$$

By the same logic as above, if the left-hand side is non-negative, at least one of the elements in the sum on the right-hand side with $p_{-i}(y_{-i}) > 0$ must be non-negative. Hence $\exists y \in X$ with $p(y) > 0$ such that $\pi_i(x_i, y_{-i}) - \pi_i(y) \geq 0$ *Q.E.D.*

Lemma 1 says that if a pure strategy x_i is a better reply to the product distribution p, then there also exists a pure strategy-tuple in the support of p to which x_i is a better reply. This implies that the set of better replies to all product distributions over a particular product set of pure strategy-tuples is identical to the set of better replies to the product set of pure strategy-tuples.

Corollary 1. $\gamma(\square(Y)) = \tilde{\gamma}(Y)$ *for all* $Y \in \mathcal{X}$.

It is worth noting that a result analogous to Lemma 1 does not hold for the best-reply correspondence. Consider the game in Figure 1, and the distribution $p^* = ((\frac{1}{2}, \frac{1}{2}, 0), (\frac{1}{2}, \frac{1}{2}))$ with the associated expected payoff $u_1(p^*) = \frac{3}{2}$ to player 1. Clearly, C is a best reply for player 1 to p_2^*, and hence also a better reply to p^*. Moreover, there exists a pure strategy-tuple (B, a) in the support of p^* to which C is a better reply. However, there exists no pure strategy in the support of p_2^* to which C is a best reply since $\beta_1^{-1}(C) = \{p_2 \in \Delta(X_2) \mid \frac{1}{3} \leq p_2(a) \leq \frac{2}{3}\}$.

	a	b
A	$3,1$	$0,0$
B	$0,0$	$3,1$
C	$2,0$	$2,0$

FIGURE 1

Recall that $H = X^m$ is the space of histories. Let H' be a subset of this set and let $S(H')$ be the *span* of this subset, i.e. the product set of all pure strategies that appear in some history in H'. H' is a *minimal configuration closed under* γ if $S(H')$ is a minimal closed set under better replies. We can now state the main theorem of the paper.

Theorem 1. *Let* Γ *be a* n*-player game on the finite strategy space* X *and let* $P^{\gamma,m,s,\varepsilon}$ *be* γ*-adaptive play. If* s/m *is sufficiently small and* s *is sufficiently large, the unperturbed process* $P^{\gamma,m,s,0}$ *converges with probability one to a minimal configuration* *closed under γ. If, in addition, ε is sufficiently small, the perturbed process $P^{\gamma,m,s,\varepsilon}$ puts arbitrarily high probability on the minimal configuration(s) closed under γ that minimizes the stochastic potential.*

PROOF: See the Appendix.

Theorem 1 resembles Theorem 7.2 in Young (1998, p. 111) but does not require the game to be nondegenerate in any sense. This simplifies the first part of the proof which in other respects is analogous to Young's proof. Young requires that the game belong to the generic class of games with the following property.

Definition 4. (Young, 1998) Γ is *nondegenerate in best replies* (NDBR) if for every i and every x_i, either $\beta_i^{-1}(x_i)$ is empty or it contains a nonempty subset that is open in the relative topology of $\square(X_{-i})$.

3.2. Other Results. We will now proceed by demonstrating that two other results for β-adaptive play in Young (1998) can be obtained for γ-adaptive play in a straightforward manner. These results follow since γ-adaptive play is a *regular perturbed Markov process* (see the Appendix) and since by the definition of the better-reply correspondence $\beta(p) \subset \gamma(p)$, and $\gamma(p) = x$ with $p(x) = 1$ if and only if x is a strict Nash equilibrium.

Young shows that a state h is an absorbing state of β-adaptive play if and only if it consists of a strict Nash equilibrium played m times in succession. He calls such a state a *convention*. Due to the two last properties mentioned above, the same relation holds also for γ-adaptive play. It is clear that an absorbing state of this process must be a state on the form $h = (x, ..., x)$. Moreover, x must be a unique better reply to a distribution p with $p(x) = 1$ and hence a strict Nash equilibrium. Conversely, any state consisting of m repetitions of a strict Nash equilibrium is clearly an absorbing state.

Theorem 2. *Let Γ be a 2×2 coordination game, and let $P^{\gamma,m,s,\varepsilon}$ be γ-adaptive play.*
(i) If information is sufficiently incomplete ($s/m \leq 1/2$), then from any initial state, the unperturbed process $P^{\gamma,m,s,0}$ converges with probability one to a convention and locks in.
(ii) If information is sufficiently incomplete ($s/m \leq 1/2$), and s and m are sufficiently large, the stochastically stable states of the perturbed process correspond one to one with the risk-dominant equilibria.

PROOF: Replace the words "best reply" and "best replies" by "better reply" and "better replies", and delete "therefore" on line 31 p. 69 in the proof of Theorem 4.2 in Young (1998, pp. 68-70). *Q.E.D.*

Given n-player game Γ with finite strategy space $X = \prod X_i$, associate each strategy-tuple $x \in X$ with the vertex of a graph. Draw a directed edge from vertex x to vertex x' if and only if: i) there exists exactly one player i such that $x_i \neq x'_i$, $\pi_i(x'_i, x_{-i}) \geq \pi_i(x)$ and ii) there does not exist $x''_i \neq x_i$ and $x''_i \neq x'_i$ such that $\pi_i(x'_i, x_{-i}) > \pi_i(x''_i, x_{-i}) > \pi_i(x)$. The graph obtained in this manner is called a *better-reply graph* (Josephson and Matros, 2000). A *better-reply path* is a sequence of the form $x^1, x^2, ..., x^l$ such that each pair (x^j, x^{j+1}) corresponds to a directed edge in the better-reply graph. We say that a game is *weakly acyclic in* γ if there from every vertex exist a directed path to a *sink*, a vertex with no outgoing edges.

Theorem 3. *Let Γ be a game weakly acyclic in γ, and let $P^{\gamma,m,s,\varepsilon}$ be γ-adaptive play. If s/m is sufficiently small, the unperturbed process $P^{\gamma,m,s,0}$ converges with probability one to a convention from any initial state. If, in addition, ε is sufficiently small, the perturbed process puts arbitrarily high probability on the conventions(s) that minimize stochastic potential.*

Proof: Replace $P^{m,s,o}$ by $P^{\gamma,m,s,o}$, and the words "best replies" and "best-reply path" by "better replies" and "better-reply path" respectively in the proof of Theorem 7.1 in Young (1998, pp. 163-164). *Q.E.D.*

Clearly if a game is weakly acyclic in β (that is weakly acyclic according to Young's (1998) definition), then it is also weakly acyclic in γ. The opposite may not hold, as illustrated by the game in Figure 4. In this game there is a better-reply path from any vertex to $\{C\} \times \{c\}$, but no best-reply path from any vertex in $\{A, B\} \times \{a, b\}$ to this unique sink.

	a	b	c
A	$2,1$	$1,2$	$0,0$
B	$0,2$	$2,1$	$0,0$
C	$1,0$	$0,0$	$2,2$

Figure 2

4. Relation to Regular Selection Dynamics

In this section we relate the above result concerning the connection between minimal sets closed under better replies and stochastically stable states under γ-adaptive play to Ritzberger's and Weibull's (1995) findings in a deterministic continuous-time model. Apart from the stochastic element in our model, there is another important difference between the approaches in these two models. Whereas we make detailed assumptions about individual behavior in the sense that all individuals are assumed to apply a specific adaptive rule, Ritzberger and Weibull (1995) make general assumptions about aggregate population dynamics. Yet, there is a clear connection between the results in the two models, as will be discussed below. We first reprint two definitions and an important theorem from Ritzberger and Weibull (1995).

Definition 5. *(Ritzberger and Weibull, 1995) A regular selection dynamics on* $\square(X)$

is a system of ordinary differential equations: $\dot{p}_i^k = f_i^k(p)p_i^k \; \forall k = 1, ..., |X_i|, \forall i \in N$ *with* $f_i : \square(X) \to R^{|X_i|}$ *and* $f = \prod_{i \in N} f_i$ *is such that*
(i) f *is Lipschitz continuous on* $\square(X)$
(ii) $f_i(p) \cdot p_i = 0 \; \forall p \in \square(X), \forall i \in N$.

Definition 6. *(Ritzberger and Weibull, 1995) A sign-preserving selection dynamics (SPS) is a regular selection dynamics such that for all* $i \in N$*, all* $p \in \square(X)$ *and all* x_i^k *such that* $p_i(x_i^k) > 0$ $u_i(x_i^k, p_{-i}) < u_i(p) \Leftrightarrow f_i^k(p) < 0$.

Theorem 4. *(Ritzberger and Weibull, 1995) For any SPS dynamics and any set* $Y \in \mathcal{X}$, $\square(Y)$ *is asymptotically stable if and only if* Y *is closed under better replies.*

We can make two general observations by comparing Theorem 1 with Theorem 4. First, our result is consistent with that for SPS dynamics in the sense that the limiting distribution of γ-adaptive play puts positive probability only on histories with support in sets which are asymptotically stable in SPS dynamics. Second, our result provides a tool for selection among different asymptotically stable sets. The unperturbed process selects those asymptotically stable sets which correspond to minimal closed sets under better replies and the perturbed process those minimal closed set(s) that minimize the stochastic potential.

In interpreting the above results it is important to keep in mind the distinction between asymptotic stability and stochastic stability. Asymptotic stability refers to robustness against small one-time shocks. It is a local property since it only assures that nearby points will converge to the stable set. Stochastic stability refers to robustness against perpetual random shocks. It is a global property in the sense that the perturbed process, independently of the initial state, in the long run will spend most of the time in the stochastically stable states.

5. Examples

Consider the 2-player game in Figure 3, taken from Young (1998). This game has two closed sets under better replies, X and $\{D\} \times \{d\}$, of which the latter is a minimal set. This set is also the unique minimal closed set under best replies. Young shows that this game is nondegenerate in best replies since strategy B is a best reply to a set of distributions which is neither empty nor contains a set open in $\square(X_2)$. The unperturbed version of β-adaptive play has two recurrent classes, one with span $\{A, C\} \times \{a, c\}$ and one with span $\{D\} \times \{d\}$. However, it is easy to see that B is a better reply to several pure strategy-tuples and by Theorem 1 γ-adaptive play converges with probability one to the state $h = ((D, d), ..., (D, d))$. By Theorem 4, $\{D\} \times \{d\}$ is asymptotically stable under SPS dynamics.

	a	b	c	d
A	$0, 1$	$0, 0$	$2, 0$	$0, 0$
B	$2/(1+\sqrt{2}), 0$	$-1, 1/2$	$2/(1+\sqrt{2}), 0$	$0, 0$
C	$2, 0$	$0, 0$	$0, 1$	$0, 0$
D	$0, 0$	$1, 0$	$0, 0$	$2, 2$

Figure 3

The game in Figure 4 is nondegenerate in β. It has four sets closed under β: X, $Y = \{B\} \times \{b\}$ and $Z = \{A, B\} \times \{a, b\}$ and $T = \{C\} \times \{c\}$. It has three sets closed under γ: X, Y and T. The two minimal closed sets under β are identical to the two minimal closed set under γ: Y and T. According to Theorem 4, $\square(X)$, $\square(Y)$ and $\square(T)$ are asymptotically stable under SPS dynamics. By Theorem 1 it follows that the unperturbed version of γ-adaptive play converges with probability one to either Y or T. Moreover, it is easy to check that T has minimum stochastic potential implying that the state $h = ((C, c), ..., (C, c))$ is stochastically stable.

	a	b	c
A	$2,5$	$2,5$	$0,0$
B	$0,0$	$3,1$	$0,0$
C	$1,3$	$0,0$	$4,4$

FIGURE 4

6. EXTENSIONS

The better-reply and the best-reply correspondences both belong to the class of *behavior correspondences* (Ritzberger and Weibull, 1995). Behavior correspondences are upper hemi-continuous correspondences such that the image of any mixed strategy profile includes the set of best-replies. We believe that the main result in this paper can be extended in a straight-forward manner to other correspondences in this class if it is assumed that the game is nondegenerate in a specific sense. In other words, our conjecture is that each behavior correspondence φ give rise to a regular perturbed Markov process and that the unperturbed version of this process converges with probability one to a minimal configuration closed under φ. If, in addition, the mutation rate is sufficiently small, the perturbed version of the process puts arbitrarily high probability on the minimal configuration(s) closed under φ that minimize stochastic potential. This result requires that the set of replies to all sample distributions over a particular set be identical to the set of replies to all possible distributions over the same set. For some correspondences, such as the best-reply correspondence, this is assured if it is assumed that the game fulfils certain conditions. For other correspondences, like the better-reply correspondence, such an assumption is not needed.

Department of Economics, Stockholm School of Economics, P.O. Box 6501, S-113 83 Stockholm, Sweden.

APPENDIX

Definition 7. (Young, 1998) P^ε is a *regular perturbed Markov process* if P^ε is irreducible for every $\varepsilon \in (0, \varepsilon^*]$, and for every $h, h' \in H$, $P^\varepsilon_{hh'}$ approaches $P^0_{hh'}$ at an exponential rate, i.e. $\lim_{\varepsilon \to 0} P^\varepsilon_{hh'} = P^0_{hh'}$ and if $P^\varepsilon_{hh'} > 0$ for some $\varepsilon > 0$, then $0 < \lim_{\varepsilon \to 0} \frac{P^\varepsilon_{hh'}}{\varepsilon^{r_{hh'}}} < \infty$ for some $r_{hh'} \geq 0$.

Lemma 2. *γ-adaptive play is a regular perturbed Markov process.*

PROOF: This proof follows Young (1998, p. 55) closely. The process $P^{\gamma,m,s,\varepsilon}$ operates on a the finite space X^m of length-m histories. Given a history $h = (x^1, x^2, ..., x^m)$ at time t, the process moves in the next period to a state of form $h' = (x^2, x^3 ..., x^m, x)$ for some $x \in X$. Recall that any such state h' is said to be a successor of h. Before choosing an action, an agent in player position i draws a sample of size s from the m previous choices in h for each population $j \in \{1, ..., n\}$ (including his own), the samples being independent among distinct populations j. The action x_i is an idiosyncratic choice or error if and only if there exists no set of n samples in h (one from each population j) such that $x_i \in \gamma_i(\hat{p})$, where $\hat{p}$ is the product of the frequency distributions. For each successor h', let $r_{hh'}$ denote the total number of errors in the rightmost element of h'. Evidently $0 < r_{hh'} < n$. It is easy to see that the probability of the transition $h \to h'$ is on the order of $\varepsilon^{r_{hh'}}(1-\varepsilon)^{n-r_{hh'}}$. If h' is not a successor of h, the probability of the transition $h \to h'$ is zero. Thus the process $P^{\gamma,m,s,\varepsilon}$ approaches $P^{\gamma,m,s,0}$ at a rate that is approximately exponential in ε; furthermore it is irreducible whenever $\varepsilon > 0$. It follows that $P^{\gamma,m,s,\varepsilon}$ is a regular perturbed Markov process. *Q.E.D.*

Theorem 5. *(Young, 1998) Let P^ε be a regular perturbed Markov process and let μ^ε be the unique stationary distribution of P^ε for $\varepsilon > 0$. Then $\lim_{\varepsilon \to 0} \mu^\varepsilon = \mu^0$ exists, and* *is a stationary distribution of P^0. The stochastically stable states are precisely those states that are contained in the recurrent classes of P^0 having minimum stochastic potential.*

Proof of Theorem 1: Except for the first part, this proof is analogous to the proof of Theorem 7.2 in Young (1998, pp. 164-166). In order to prove the first claim, we show that for sufficiently large s and sufficiently small s/m, the spans of the recurrent classes of the unperturbed process correspond one to one with the minimal sets closed under γ of the game.

First note that by Corollary 1 $\gamma_i(\Box(Y)) = \tilde{\gamma}_i(Y)$ for all $Y \in \mathcal{X}$. This implies that the set of better replies to the distributions on an arbitrary set of strategy-tuples is identical to the set of better replies to the sample distributions on the same set.

For each product set Y and player i, define the mapping $\check{\gamma}_i(Y) = Y_i \cup \gamma_i(\Box(Y))$ and $\check{\gamma}(Y) = \prod \check{\gamma}_i(Y)$. Choose m such that $m \geq s\,|X|$. Furthermore, fix a recurrent class E_k of $P^{\gamma,m,s,0}$, and choose any $h^0 \in E_k$ as the initial state. We shall show that the span of E_k, $S(E_k)$, is a minimal closed set under γ. It is evident that there is a positive probability of reaching a state h^1 in which the most recent s entries involve a repetition of some fixed $x^* \in X$. Note that $h^1 \in E_k$, because E_k is a recurrent class. Let $\check{\gamma}^{(j)}$ denote the j-fold iteration of $\check{\gamma}$ and consider the nested sequence:

$$\{x^*\} \subseteq \check{\gamma}(\{x^*\}) \subseteq \check{\gamma}^{(2)}(\{x^*\}) \subseteq ... \subseteq \check{\gamma}^{(j)}(\{x^*\}) \subseteq ... \tag{3}$$

Since X is finite, there exists some point at which this sequence becomes constant, say, $\check{\gamma}^{(j)}(\{x^*\}) = \check{\gamma}^{(j+1)}(\{x^*\}) = Y^*$. By construction, Y^* is a set closed under γ.

Assume $\check{\gamma}(\{x^*\}) \neq \{x^*\}$ (otherwise the following argument is redundant). Then there is a positive probability that, beginning after the history h^1, some $x^1 \in \check{\gamma}(\{x^*\})\backslash\{x^*\}$ will be chosen for the next s periods. Call the resulting history h^2. Then there is a positive probability that $x^2 \in \check{\gamma}(\{x^*\})\backslash\{x^*, x^1\}$ will be chosen for the next s periods and so forth. Continuing in this way one eventually obtains a history h^k such that all elements of $\check{\gamma}(\{x^*\})$, including the original $\{x^*\}$, appear at least s times. All that is needed to assume is that m is large enough so that the original s repetitions of x^* have not been forgotten. This is assured if $m \geq s\,|X|$. Continuing this argument it is clear that there is a positive probability of eventually obtaining a history h^* in which all members of Y^* appear at least s times within the last $s\,|Y|$ periods. In particular, $S(h^*)$ contains Y^* which by construction is a set closed under γ.

We claim that Y^* is a minimal closed set under γ. Let Z^* be a minimal closed set under γ contained in Y^*, and choose $z^* \in Z^*$. Beginning with the history h^* already constructed, there is a positive probability that z^* will be chosen for the next s periods. After this, there is a positive probability that only elements of $\check{\gamma}(\{z^*\})$ will be chosen, or members of $\check{\gamma}^2(\{x^*\})$, or members of $\check{\gamma}^3(\{x^*\})$, and so on. This happens if agents always draw samples from the new part of the history that follows h^*, which they will do with positive probability.

The sequence $\check{\gamma}^{(k)}(\{z^*\})$ eventually becomes constant with value Z^*, because Z^* is a minimal closed set under γ. Moreover, the part of the history before the s-fold repetition of x^* will be forgotten within m periods. Thus there is a positive probability of obtaining a history h^{**} such that $S(h^{**}) \subseteq Z^*$. From such a history the process $P^{m,s,0}$ can never generate a history with members that are not in Z^* because Z^* is a set closed under γ.

Since the chain of events that led to h^{**} began with a state in E_k, which is a recurrent class, h^{**} is also in E_k; moreover, every state in E_k is reachable from h^{**}. It follows that $Y^* \subseteq S(E_k) \subseteq Z^*$, from which it can be concluded that $Y^* = S(E_k) = Z^*$.

Conversely, we must show that if Y^* is a minimal closed set under γ, then $Y^* = S(E_k)$ for some recurrent class E_k of $P^{\gamma,m,s,0}$. Choose an initial history h^0 that involves only strategies in Y^*. Starting at h^0, the process $P^{\gamma,m,s,0}$ generates histories that involve no strategies that lie outside of $S(h^0)$, $\check{\gamma}(S(h^0))$, $\check{\gamma}^{(2)}(S(h^0))$ and so on. Since Y^* is a set closed under γ, all of these strategies must occur in Y^*. With probability one the process eventually enters a recurrent class, say E_k. It follows that $S(E_k) \subseteq Y^*$. Since Y^* is a minimal closed set under γ, the earlier part of the argument shows that $S(E_k) = Y^*$. This establishes the one to one correspondence between minimal sets closed under γ and the recurrent classes of $P^{\gamma,m,s,0}$.

The second claim of Theorem 1 now follows from the fact that γ-adaptive play, by Lemma 2, is a regular perturbed Markov process, and by Theorem 5 which states that the stochastically stable states of such a process are the states that are contained in the recurrent classes of the unperturbed process with the minimum stochastic potential.

References

[1] Josephson, J., and A. Matros (2000): "Stochastic Imitation in Finite Games," SSE/EFI Working Paper Series in Economics and Finance No. 363, Stockholm.

[2] Ritzberger, K., and J. W. Weibull (1995): "Evolutionary Selection in Normal-Form Games," *Econometrica* 63, 1371-1400.

[3] Young, P. H. (1993): "The Evolution of Conventions," *Econometrica* 61, 57-84.

[4] ——— (1998): *Individual Strategy and Social Structure.* Princeton University Press.